Breaking *One-Size-Fits-All*: Revisiting Out-of-Distribution Detection on Graphs Under Diverse Distribution Shifts

Chuancheng Song[1,2], **Hanyang Shen**[1,2], **Yan Dong**[1,2], **Xixun Lin**[1,2], **Yanmin Shang**[1,2*], **Yanan Cao**[1,2]

[1]Institute of Information Engineering, Chinese Academy of Sciences
[2]School of Cyber Security, University of Chinese Academy of Sciences
{songchuancheng, shenhanyang, dongyan, linxixun, shangyanmin, caoyanan}@iie.ac.cn

Abstract

Graph OOD detection is crucial in open-world scenarios, where OOD samples may manifest in diverse forms such as *open-set deviations*, *feature-similar shifts*, and *structural anomalies*, each exhibiting distinct geometric characteristics. However, most existing methods adopt a *one-size-fits-all* geometric assumption (typically Euclidean space), which inadequately captures the diverse nature of real-world distribution shifts. Therefore, adaptively selecting geometric spaces according to the properties of OOD samples is critical for their effective representation and reliable identification. Motivated by this, we revisit the graph OOD detection task under diverse distribution shifts and propose UniGOD, a unified framework serving as a graph foundation model for this task. UniGOD comprises two core modules: GeoUP and DynEVO. GeoUP module adaptively perceives the geometric space (such as Euclidean, hyperbolic, and hyperspherical space) by learning the curvature κ of Riemannian manifolds. DynEVO module leverages the dynamic nature of neural SDEs to reveal pronounced uncertainty differences between ID/OOD samples, which are reflected in the divergent evolutionary trajectories of node embeddings induced by κ-GNN iterations. With the geometry-dynamics coupling mechanism of the above two modules, UniGOD effectively captures the diverse distribution shifts. Extensive experiments demonstrate its superior performance over existing SOTA methods.

Introduction

Graph out-of-distribution (OOD) detection plays a vital role in open-world scenarios and is predominantly approached through Graph Neural Networks (GNNs) (Dong et al. 2025b). Existing studies have explored the graph OOD detection task from various perspectives (Cai et al. 2025; Zhang et al. 2024), with the fundamental aim of learning expressive representations that can accurately characterize OOD samples and separate them from in-distribution (ID) samples. However, a common limitation among them is the implicit assumption of a fixed representation space (typically Euclidean space), which inadequately captures the diverse OOD samples in real-world distribution shifts.

Specifically, this *one-size-fits-all* geometric assumption overlooks a key observation: unlike ID data that follow the training distribution, OOD data deviate in various forms, such as *open-set deviations*, *feature-similar shifts*, and *structural anomalies*, each exhibiting distinct geometric characteristics. Fig. 1 presents illustrative case studies: (a) We use Twitch-EN as ID, and treat Twitch-FR and Twitch-ES as OOD. While Euclidean space can roughly separate the distant *open-set deviations* (Twitch-ES) with larger intervals, it struggles to distinguish *feature-similar shifts* (Twitch-FR) with closer distances. In contrast, hyperspherical space, which emphasizes angular margins, achieves clearer separation, as confirmed by visualization and angular distance metrics (e.g., $\delta(\cdot) > d(\cdot)$). (b) OOD nodes (music) are structurally misclassified as ID (camera) via strong co-purchase links (*structural anomalies*), making them indistinguishable from ID in Euclidean space. However, hyperbolic space, being more sensitive to structural variations, reveals clearer separation. Furthermore, real-world OOD often involves mixtures of these shift types, rendering *one-size-fits-all* geometry representations inherently limited. This raises a fundamental question: ***How can we adaptively select geomet-***

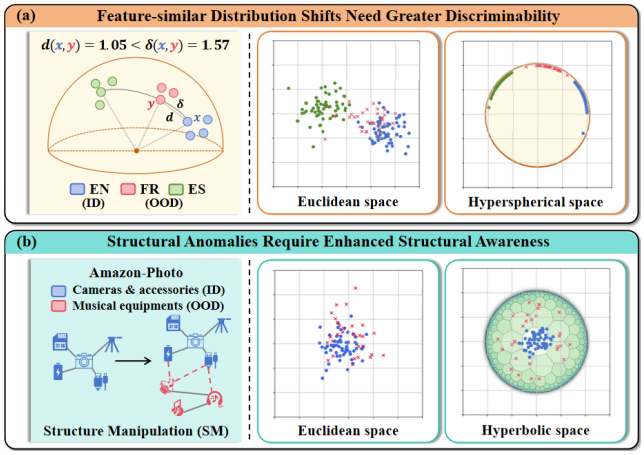

Figure 1: (a) `Twitch` is a language-partitioned social network, where the English Community lies close to Français Comm. due to shared cultural context and platform behavior, but is distant from Español Comm.; (b) `Amazon-Photo` is a co-purchase network where camera and music items may be linked via shared accessories (e.g., battery, charger).

*Corresponding author.

ric spaces according to the properties of OOD samples for their effective representation and reliable identification?

Motivated by this, we revisit the graph OOD detection task under diverse distribution shifts and identify two core challenges: **(1) Select appropriate geometric spaces for different OOD types**. On the one hand, different distribution shifts exhibit different geometric characteristics; on the other hand, some OOD samples may simultaneously involve multiple types of shifts, necessitating a superposition of geometric spaces. These highlight the need for a unified and adaptive framework that can adaptively select or integrate suitable geometric spaces according to the properties of OOD samples. **(2) Achieve reliable separation in the absence of OOD labels**. The inherently unsupervised nature of OOD detection, where no labeled OOD samples are available during training, inevitably leads to representation entanglement and boundary ambiguity between ID and OOD samples, rendering vanilla distance metrics insufficiently reliable for detecting OOD samples. This calls for a separation criterion that measures more essential differences between OOD and ID samples.

To tackle these challenges, we propose **UniGOD**, a **Uni**fied framework for **G**raph **O**ut-of-distribution **D**etection based on a novel **G**eometry–**D**ynamics coupling mechanism grounded in Riemannian manifolds. UniGOD comprises two core modules: GeoUP and DynEVO. The GeoUP module adaptively perceives the geometric space (such as Euclidean, hyperbolic, and hyperspherical space) by learning the curvature κ of Riemannian manifolds. The DynEVO module leverages the dynamic nature of neural stochastic differential equations (SDEs) to reveal pronounced uncertainty differences between ID and OOD samples, which are reflected in the divergent evolutionary trajectories of node embeddings induced by κ-GNN iterations. With the geometry-dynamics coupling mechanism of the above two modules, UniGOD effectively captures the diverse distribution shifts. In summary, our contributions are threefold:

- To the best of our knowledge, we present the first unified framework addressing diverse graph OOD shifts, serving as an initial graph foundation model tailored for the graph OOD detection task.
- We design a novel geometry–dynamics coupling mechanism grounded in manifolds that integrates curvature-adaptive embedding and trajectory-based uncertainty.
- Experiments demonstrate that UniGOD outperforms SOTA baselines across multiple OOD scenarios.

Related Work

Graph Out-of-Distribution Detection

Existing graph OOD methods mainly fall into output-based and feature-based (including distance- and uncertainty-aware metrics) methods. Output-based methods distinguish OOD/ID samples based on final predictions such as softmax, logits, or energy values (Hendrycks et al. 2017; Liang et al. 2018; Wu et al. 2023; Gong et al. 2024). Feature-based (distance-aware) methods operate on intermediate embeddings and assume OOD diverges from ID. TopoOOD (You et al. 2022) reshapes node neighborhoods to amplify distance discrepancies in latent space. GOOD-D (Liu et al. 2023) and ODGNN (Zhang et al. 2025) improve OOD/ID separability by enlarging latent feature distances through contrastive learning. Feature-based (uncertainty-aware) methods rely on representation instability to signal unfamiliarity with OOD data. GKDE (Zhao et al. 2020) uses kernel density estimation in the embedding space. GraphDE (Li et al. 2022) detects OOD samples via negative log-likelihood as the uncertainty-aware score. EviSEC (Sun et al. 2025) adopts Dirichlet evidential learning to model prediction uncertainty. Feature-based methods leverage latent representations but are confined to a single Euclidean space, which may be insufficient for diverse OOD shifts and motivates our geometry-aware exploration in this work.

Geometric Graph Representation Learning

Vanilla graph representation learning (GRL) operates in Euclidean space (Perozzi, Al-Rfou, and Skiena 2014; Grover and Leskovec 2016; Kipf and Welling 2017). However, Euclidean geometry struggles to model intricate structures and suffers from over-squashing and over-smoothing (Zhu et al. 2020; Yang et al. 2022; Zheng et al. 2025; Rusch, Bronstein, and Mishra 2023). Therefore, recent studies explore Riemannian manifolds for more expressive embeddings: (1) Hyperbolic space is structure-aware and ideal for hierarchies and exponential growth (Nickel and Kiela 2017, 2018). HGCN (Chami et al. 2019) leverages hyperbolic geometry with Poincaré and Lorentz embeddings to capture hierarchical structures for improved node classification. HBP (Song et al. 2024) improves hyperbolic GRL by introducing second-order pooling technology, enhancing the ability to capture hierarchical information. (2) Hyperspherical space, with its compactness and angular geometry, facilitates discriminative clustering (Ming et al. 2023). vMF (Wang et al. 2024) project nodes onto the hypersphere to enhance graph clustering. HyperspherE (Dong et al. 2021) and SpherE (Li, Ao, and He 2024) leverage this geometric structure in knowledge graph modeling scenarios. Despite advances in GRL, OOD detection remains underexplored. The diverse nature of OOD on graphs necessitates corresponding spaces for more expressive representation, and motivates a rethinking of the graph OOD detection task.

Preliminary

Problem Definition

Given a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{X})$, where $\mathcal{V}$ denotes the set of N nodes, $\mathcal{E}$ denotes the edge set, and $\mathbf{X} \in \mathbb{R}^{N \times F}$ is the node feature matrix, our goal is to learn a scoring function $s = \psi(\cdot)$ for each node feature $\mathbf{x}_i$. A larger s value indicates a higher likelihood of v_i belonging to $\mathcal{V}_{\text{OOD}}$. Associated with $\psi(\cdot)$, the binary decision function $\Delta(\cdot)$ is formulated as

$$\Delta(\mathbf{x}_i; \psi, \tau) = \begin{cases} v_i \in \mathcal{V}_{\text{OOD}}, & s_i > \tau, \\ v_i \in \mathcal{V}_{\text{ID}}, & s_i \leq \tau, \end{cases} \tag{1}$$

where $\tau = \mu + \xi\sigma$ is a decision threshold determined from the ID score distribution and $\mathcal{V} = \mathcal{V}_{\text{ID}} \cup \mathcal{V}_{\text{OOD}}$.

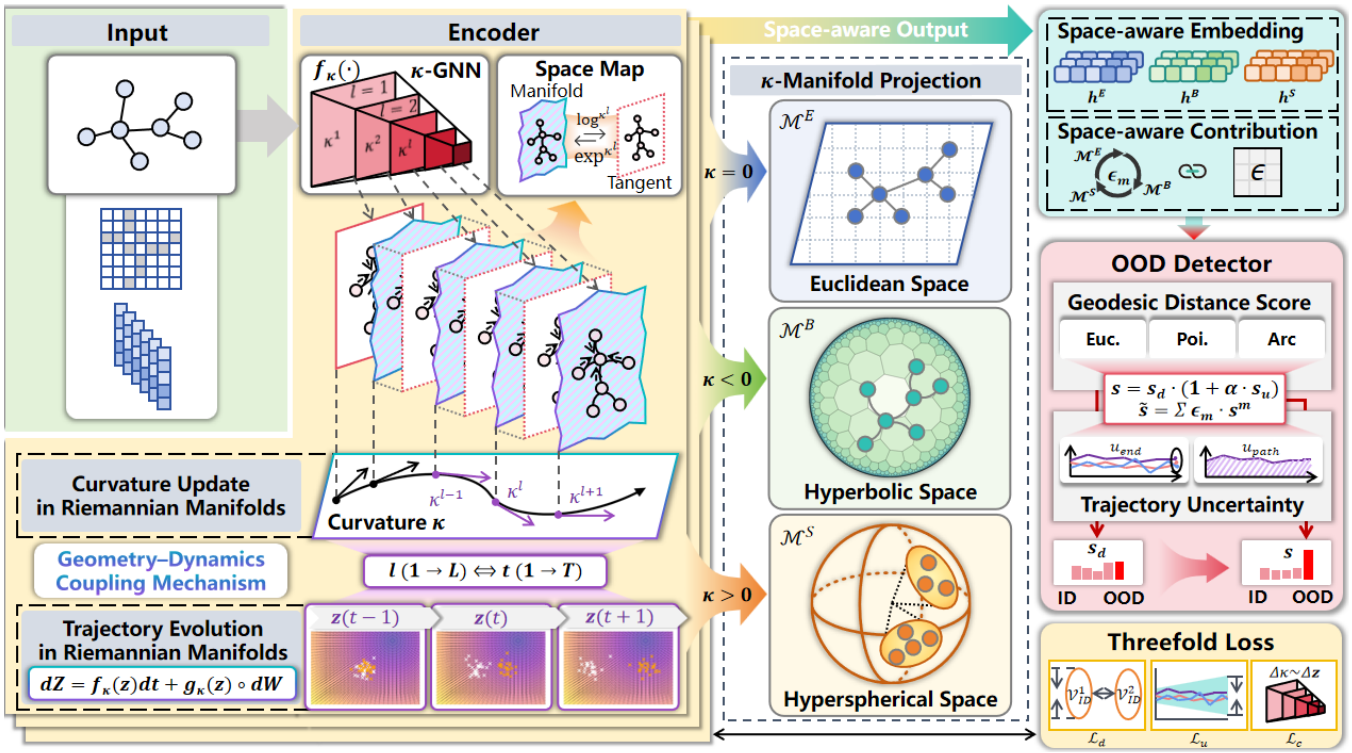

Figure 2: Overview of UniGOD: Given an input graph, UniGOD first employs a curvature-adaptive encoder (κ-GNN) to project node representations into multiple geometric spaces (GeoUP). In parallel, a neural SDE models the trajectory evolution of embeddings to quantify dynamic uncertainty with κ-GNN iterations (DynEVO). After training guided by a threefold loss, we compute a space-aware geodesic distance score and enhance it with trajectory-based uncertainty to yield the final OOD score.

Riemannian Manifolds with Constant Curvature

A constant curvature Riemannian manifold $\mathcal{M}_\kappa$ is a smooth and continuous space characterized by a constant sectional curvature κ (Sun et al. 2023). Depending on the sign of κ:

$$\mathcal{M}_\kappa = \begin{cases} \text{Hyperbolic space } \mathcal{M}_B, & \kappa < 0, \\ \text{Euclidean space } \mathcal{M}_E, & \kappa = 0, \\ \text{Hyperspherical space } \mathcal{M}_S, & \kappa > 0, \end{cases} \tag{2}$$

which provides a flexible geometric foundation to embed complex graph data with non-Euclidean structures.

Uncertainty Modeling via Neural SDEs

Neural stochastic differential equations model epistemic uncertainty with deterministic drift and stochastic diffusion:

$$d\mathbf{x}(t) = f(\mathbf{x}(t), t)\, dt + g(\mathbf{x}(t), t) \circ d\mathbf{W}(t), \tag{3}$$

where $f(\cdot)$ is the deterministic drift term, $g(\cdot)$ is the stochastic diffusion term, and $\mathbf{W}(t)$ denotes a Wiener process. The Stratonovich integral $\circ$ ensures compatibility with manifold calculus (Look et al. 2022). This formulation allows us to simulate the difference in trajectory evolution of embeddings between ID and OOD representations.

Methodology

Geometric Embedding with Curvature Update

GeoUP unifies manifolds $\mathcal{M}_\kappa$ via learnable curvature κ, enabling the optimal embedding space for more expressive representations with geometric GNN (abbr. κ-GNN).

κ-GNN In vanilla GNNs, node representations are updated via linear neighbor aggregation in Euclidean space. In contrast, κ-GNN $f_\kappa(\cdot;\theta)$ generalizes this process to non-Euclidean spaces, allowing each layer l to operate on distinct manifolds with different curvatures $\kappa^{(l)}$, thus requiring more sophisticated log-exp mappings (Dong et al. 2025a). Specifically, node i obtains an embedding $\mathbf{h}_i \in \mathcal{M}_\kappa$, updated as:

$$\mathbf{h}_i^{(l+1)} = \exp_{\mathbf{b}}^{\kappa^{(l)}}\Big(\sum\nolimits_{j\in\mathcal{N}_i} w_{ij} \cdot \log_{\mathbf{b}}^{\kappa^{(l)}}(\mathbf{h}_j^{(l)})\Big), \tag{4}$$

which can be iteratively mapped between tangent space and manifold space, ensuring effective aggregation in non-Euclidean geometries. Specifically, the logarithmic map $\log_{\mathbf{b}}^{\kappa}(\cdot) : \mathcal{M}_\kappa \to T_{\mathbf{b}}\mathcal{M}_\kappa$ projects a point from the manifold onto the tangent space at a specified *base point* $\mathbf{b}$ (typically set as $\mathbf{h}_i^{(l)}$ in practice):

$$\log_{\mathbf{b}}^{\kappa}(\mathbf{x}) = \frac{\delta}{\mathrm{sn}_\kappa(\delta)}\,(\mathbf{x} - \mathrm{cs}_\kappa(\delta)\,\mathbf{x}), \tag{5}$$

where $\delta = \delta_\kappa(\mathbf{b}, \mathbf{x})$ denotes the geodesic distance between $\mathbf{b}$ and $\mathbf{x}$ on $\mathcal{M}_\kappa$. The exponential map $\exp_{\mathbf{b}}^{\kappa}(\cdot) : T_{\mathbf{b}}\mathcal{M}_\kappa \to$

$\mathcal{M}_\kappa$ then maps a tangent vector back onto the manifold:

$$\exp_{\mathbf{b}}^{\kappa}(\mathbf{y}) = \mathrm{cs}_\kappa(\|\mathbf{y}\|)\,\mathbf{b} + \mathrm{sn}_\kappa(\|\mathbf{y}\|)\,\frac{\mathbf{y}}{\|\mathbf{y}\|}, \tag{6}$$

where $\mathbf{y}$ is on $T_{\mathbf{b}}\mathcal{M}_\kappa$; $\mathrm{cs}_\kappa(\cdot)$ and $\mathrm{sn}_\kappa(\cdot)$ are the generalized cosine and sine functions for providing a unified trigonometric formulation across different geometric spaces:

$$\mathrm{cs}_\kappa(x), \mathrm{sn}_\kappa(x) = \begin{cases} \cos(\sqrt{\kappa}x), \frac{\sin(\sqrt{\kappa}x)}{\sqrt{\kappa}}, & \kappa > 0, \\ 1, x, & \kappa = 0, \\ \cosh(\sqrt{-\kappa}x), \frac{\sinh(\sqrt{-\kappa}x)}{\sqrt{-\kappa}}, & \kappa < 0. \end{cases} \tag{7}$$

With iterative GNN propagation, κ is progressively updated to refine node embedding on the corresponding manifold $\mathcal{M}_\kappa$. Once $\kappa^{(l)}$ converges to minimize Eq. 19, the resulting geometry facilitates clear separation of ID classes and gets the optimal embedding $\mathbf{h}_i^{\kappa^*}$ and provides a natural foundation for measuring distributional deviations via geodesic distance (Crane et al. 2020).

Geodesic Distance As mentioned above, geodesic distance $\delta_\kappa(\mathbf{x}, \mathbf{y})$ depends on the curvature κ and the chosen manifold representation. In our framework, we adopt curvature-aware distances defined as:

$$\delta_\kappa(\mathbf{x}, \mathbf{y}) = \begin{cases} \frac{1}{\sqrt{\kappa}} \arccos(\kappa\langle \mathbf{x}, \mathbf{y}\rangle), & \kappa > 0, \\ \|\mathbf{x} - \mathbf{y}\|, & \kappa = 0, \\ \frac{1}{\sqrt{-\kappa}} \operatorname{arccosh}(1 - \kappa\langle \mathbf{x}, \mathbf{y}\rangle), & \kappa < 0, \end{cases} \tag{8}$$

where $\langle \mathbf{x}, \mathbf{y}\rangle$ denotes the inner product in the ambient space. Geodesic distance δ_κ seamlessly corresponds to different distributional modeling paradigms, ensuring distance measurements remain consistent and geometrically meaningful.

To quantify the deviation of a test node from the ID regions, we introduce a geodesic distance-based score. Specifically, after obtaining the final node embedding $\mathbf{h}_i$, we compute its minimum geodesic distance to all ID class centers $\{\mu_c\}_{c=1}^{C}$ (defined as the Fréchet mean):

$$s_d(i) = \min_c \delta_\kappa(\mathbf{h}_i, \mu_c), \tag{9}$$

where $\delta_\kappa(\cdot, \cdot)$ denotes the geodesic distance on $\mathcal{M}_\kappa$. A larger $s_d(i)$ indicates a higher likelihood of being OOD.

Dynamic Uncertainty with Trajectory Evolution

While GeoUP offers geodesic distances, DynEVO models uncertainty through dynamic stochastic trajectories. Inspired by Neural SDEs (Bergna et al. 2025), it captures the evolution of embeddings across κ-GNN layers, where OOD nodes exhibit higher uncertainty, reflected by more dispersed trajectories, while ID nodes show lower uncertainty with more consistent paths. These cues are learned in an unsupervised manner and complement the geometric signal.

κ-GNN SDE The stochastic trajectory is described by κ-GNN SDE, consisting of drift function f_κ and diffusion function g_κ. Following the *time-layer* equivalence $t \leftrightarrow l$, node features iterate with the discrete update rule Eq. 4:

$$\mathbf{H}_i(t+1) = \mathbf{H}_i(t) + f_\kappa(\mathbf{h}_i(t)), \tag{10}$$

where $\mathbf{H}_i(t) = \log_{\mathbf{b}}^{\kappa^{(t)}}(\mathbf{h}_i(t))$ for legal Euclidean addition. Considering a small time step Δt, we have:

$$\mathbf{H}_i(t + \Delta t) = \mathbf{H}_i(t) + \Delta t \cdot f_\kappa(\mathbf{h}_i(t)), \tag{11}$$

and in the limit as $\Delta t \to 0$, the discrete GNN update can be viewed as a first-order Euler step of an underlying ODE:

$$d\mathbf{H}_i(t) = f_\kappa(\mathbf{h}_i(t))\,dt, \tag{12}$$

where f_κ is the geometry-aware aggregation function parameterized by a κ-GNN. To further model the stochastic evolution of node trajectories, we extend Eq. 12 with Ornstein-Uhlenbeck process (denote $\mathbf{h}_i(t)$ as $\mathbf{z}_i(t)$ to distinguish discrete-layer embeddings from continuous trajectory representations):

$$d\mathbf{Z}_i(t) = f_\kappa(\mathbf{z}_i(t))\,dt + g_\kappa(\mathbf{z}_i(t)) \circ d\mathbf{W}_i(t), \tag{13}$$

where g_κ controls the stochastic diffusion. It enables stochastic trajectories evolution of node representations, forming the basis for uncertainty quantification.

Trajectory Uncertainty With Monte Carlo sampling of N stochastic trajectories from the κ-GNN SDE, we obtain their terminal embeddings $\{\mathbf{h}_i^{(n)}(T)\}_{n=1}^{N}$. To comprehensively characterize trajectory uncertainty, we consider both final-state uncertainty u_{end} and path-wise uncertainty u_{path}. The final-state uncertainty is defined as

$$u_{\text{end}}(i) = \frac{1}{N}\sum_{n=1}^{N} \delta_\kappa^2\big(\bar{\mathbf{z}}_i, \mathbf{z}_i^{(n)}(T)\big), \tag{14}$$

where $\bar{\mathbf{z}}_i$ denotes the Fréchet mean of the terminal points at time T. While relying solely on the final state may discard intermediate dynamic information, we further capture the path-wise uncertainty along the entire trajectory:

$$u_{\text{path}}(i) = \frac{1}{T}\int_0^T \frac{1}{N}\sum_{n=1}^{N} \delta_\kappa^2\big(\bar{\mathbf{z}}_i(t), \mathbf{z}_i^{(n)}(t)\big)dt, \tag{15}$$

where $\bar{\mathbf{z}}_i(t)$ represents the Fréchet mean at each intermediate time t. Finally, the unified uncertainty score $s_u(i) \in [0, 1]$ is defined as

$$s_u(i) = \beta \cdot \tilde{u}_{\text{path}}(i) + (1 - \beta) \cdot \tilde{u}_{\text{end}}(i), \tag{16}$$

where $\tilde{u}$ denotes the normalized form, and $\beta \in [0, 1]$ balances them. $s_u(i)$ explicitly characterizes the dynamic stability of node representations and their sensitivity to geometric perturbations, providing deeper insights into the intrinsic behavior of ID and OOD samples.

Unified OOD Detection Score

Intuitively, the distance score $s_d(i)$ serves as the fundamental basis for OOD detection, while the uncertainty score $s_u(i)$ acts as an amplification factor to emphasize distributional deviations. The overall OOD score is formulated as

$$s(i) = s_d(i) \cdot \big(1 + \alpha \cdot s_u(i)\big), \tag{17}$$

where α controls the uncertainty amplification effect. This combination effectively enlarges the margin between OOD and ID samples, thereby improving discriminability.

		Cora			Photo			Twitch			YelpChi	Fraud	Avg.
		LL	FI	SM	LL	FI	SM	LL	FI	SM	Real-world anomaly		
Output-based methods	MSP	0.7612	0.7587	0.7515	0.8543	0.8706	0.8689	0.6610	0.6832	0.3368	0.5715	0.8935	0.7464
	ODIN	0.3068	0.2122	0.2035	0.1027	0.0537	0.0351	0.4383	0.4227	0.3072	0.5831	0.8988	0.3687
	Energy	0.8325	0.7933	0.6958	0.8717	0.8454	0.8096	0.5842	0.5435	0.5431	0.4938	0.9033	0.7369
	GNNSAFE	0.9237	0.8924	0.8206	0.9333	0.9191	0.9064	0.5103	0.5215	0.5218	0.5533	0.9627	0.7908
	EnergyDef	0.9426	0.8935	0.8552	0.9309	0.9221	0.8883	0.6640	0.5982	0.5981	0.6205	0.8658	0.8137
Feature-based methods	Mahalanobis	0.6751	0.6391	0.4109	0.6246	0.7244	0.6083	0.4679	0.4109	0.4108	0.5320	0.7559	0.5627
	KNN	0.8464	0.8081	0.7951	0.7636	0.5452	0.6293	0.5955	0.5783	0.4980	0.5653	0.8477	0.6795
	GKDE	0.9179	0.8807	0.8221	0.8992	0.8648	0.9034	0.8015	0.8621	0.8226	0.5815	0.7732	0.7955
	GPN	0.8363	0.5703	0.5375	0.8919	0.7093	0.7147	0.7742	0.5377	0.5374	0.6825	0.8263	0.7192
	GraphDE	**0.9563**	0.7802	0.5380	0.8501	0.7906	0.7238	0.6396	0.6006	0.5381	0.5603	0.6701	0.6654
	ODGNN	0.8207	0.7641	0.8660	0.8674	0.7272	0.8407	0.7207	0.7144	0.8153	0.6641	0.9122	0.7936
	TopoOOD	0.9350	0.6956	0.7955	0.9410	0.8942	0.9441	0.8082	0.7877	0.8284	0.6650	0.9668	0.8267
Geometric GRL methods*	HGCN	0.8519	0.6825	0.8836	0.9201	0.8912	0.9407	0.8059	0.7563	0.8251	0.5551	0.7799	0.8299
	HYLA	0.8633	0.8118	0.9037	0.8801	0.8226	0.9273	0.7438	0.8307	0.8224	0.5505	0.7192	0.7825
	CIDER	0.8801	0.9032	0.8176	0.9407	**0.9751**	0.6377	0.8145	0.8417	0.7726	0.5537	0.7072	0.8063
	SpaceGNN	0.9105	0.9243	0.9007	0.9243	0.9201	0.9346	0.8805	0.8513	0.8435	0.6566	0.9331	0.8713
	UniGOD	0.9413	**0.9256**	**0.9244**	**0.9702**	0.9519	**0.9561**	**0.8948**	**0.8826**	**0.9113**	**0.6961**	**0.9691**	**0.9016**

Table 1: OOD detection measured by AUROC (↑) with three OOD types (Label Leave-out, Feature Interpolation, Structure Manipulation) for synthetic OOD datasets (Cora, Amazon-Photo, Twitch) and real-world anomaly datasets (YelpChi, Amazon-Fraud). Best and second-best results are boldfaced and underlined.

Furthermore, $s(i)$ is computed within the optimal manifold, maximizing ID class separability to implicitly guide the OOD identification. In practice, OOD samples may exhibit mixed distributional patterns, rendering the superposition of their geometric nature. Therefore, we further constrain the curvature κ to distinct ranges ($\kappa < 0$, $\kappa = 0$, and $\kappa > 0$), and iteratively perform κ-GNN within each space to obtain their respective optimal embeddings and scores $s^m(i)$ to fuse them:

$$\tilde{s}(i) = \sum_m \tilde{\epsilon}_m \cdot s^m(i),\ m \in \{\mathcal{M}^E, \mathcal{M}^B, \mathcal{M}^S\}, \quad (18)$$

where $\tilde{\epsilon}_m$ denotes the fusion weight for manifold m defined by normalized Generalized Fisher Ratio.

Training Objective

In GeoUP, enhancing the intra-class compactness and inter-class separability of ID samples is crucial. To explicitly encourage this geometric discriminability, we generalize Fisher Ratio ϵ as follows:

$$\mathcal{L}_d = -\epsilon = -\frac{\sum_{c=1}^{C} \delta_\kappa^2(\mu_c, \mu)}{\sum_{c=1}^{C} \frac{1}{N_c} \sum_{i=1}^{N_c} \delta_\kappa^2(\mathbf{h}_i^{(c)}, \mu_c)}, \quad (19)$$

where $\{\mu_c\}_{c=1}^C$ are ID class centers (defined as the Fréchet mean of class c) and μ is the global center among them.

In DynEVO, we encourage the trajectory stability and low uncertainty of ID samples. Thus, the final-state variance and trajectory path integral should be minimized:

$$\mathcal{L}_u = \frac{1}{N_{\text{ID}}} \sum_{i \in \mathcal{V}_{\text{ID}}} \Big(\beta \cdot \tilde{u}_{\text{path}}(i) + (1-\beta) \cdot \tilde{u}_{\text{end}}(i) \Big). \quad (20)$$

To further ensure geometric stability and encourage consistent evolution of node representations, we introduce a geometry-dynamics coupling loss:

$$\mathcal{L}_c = \frac{1}{L} \sum_{l=0}^{L} (\Delta\kappa) \cdot \Big(\frac{1}{N} \sum_{i \in \mathcal{V}_{\text{ID}}} \Delta \mathbf{z}_i \Big), \quad (21)$$

where $\Delta\kappa = |\kappa^{(l+1)} - \kappa^{(l)}|$ denotes the curvature change between consecutive layers, and $\Delta\mathbf{z}_i = \delta_\kappa(\mathbf{z}_i(l+1), \mathbf{z}_i(l))$ measures the geodesic displacement of node i. In summary, the **overall threefold loss** function is $\boldsymbol{\mathcal{L}} = \boldsymbol{\lambda}_1 \boldsymbol{\mathcal{L}}_d + \boldsymbol{\lambda}_2 \boldsymbol{\mathcal{L}}_u + \boldsymbol{\lambda}_3 \boldsymbol{\mathcal{L}}_c$, where $\lambda_1, \lambda_2, \lambda_3$ are balance parameters. Representations learned in the training guided by this threefold loss will then be fed into the OOD detector for inference.

Experiments

We would answer the following research questions:

- **RQ1**: Can UniGOD handle the diverse OOD types, outperforming SOTA methods at each type?
- **RQ2**: How does the geometry-dynamics coupling mechanism contribute to overall OOD detection performance?
- **RQ3**: Can adaptive curvature identify optimal geometric spaces for different OOD types?
- **RQ4**: How do key parameters influence the experiments?
- **RQ5**: Can UniGOD provide interpretable insights for its OOD detection decisions?

Experiment Setup

Datasets Following (Chen et al. 2025), we conduct experiments on four synthetic OOD datasets (Ogbn-ArXiv, Cora, Amazon-Photo, and Twitch) and two real-world anomaly datasets (YelpChi and Amazon-Fraud). They span citation networks, product graphs, social networks, and review networks, including three OOD types. Specifically, we use **Label Leave-out (LL)** to simulate unseen classes

(*open-set deviations*), **Feature Interpolation (FI)** to generate subtle feature perturbations close to ID clusters (*feature-similar shifts*), and **Structure Manipulation (SM)** to introduce topological alterations (*structural anomalies*). Our code is available at https://github.com/ttbook-1/UniGOD.

Baselines Our model is benchmarked against three categories of SOTA baseline methods: output-based methods (`MSP`, `ODIN`, `Energy`, `GNNSAFE`, `EnergyDef`), feature-based Euclidean methods (`Mahalanobis`, `KNN`, `GKDE`, `GPN`, `GraphDE`, `ODGNN`, `TopoOOD`), and geometric GRL methods* (`HGCN`, `HYLA`, `CIDER`, `SpaceGNN`). Methods* are originally designed for representation learning rather than OOD detection. We adapt them using standard GNN backbone and simple post-hoc scoring for fair comparison.

Performance Comparison (RQ1)

Table 1 & Fig. 3 collectively reveal four key observations: **(1) Superior overall performance of UniGOD**: UniGOD achieves the highest AUROC across all OOD types, demonstrating robust performance particularly on SM and FI, and maintaining SOTA results even on real-world anomaly datasets. **(2) Limitations of Euclidean baselines**: Output and feature-based methods in Euclidean space show relatively strong performance on LL, but struggle on SM and show suboptimal performance on FI. This highlights their insufficient geometric expressiveness and limited capacity to accommodate diverse OOD types. **(3) Restricted generalizability of single Riemannian space**: While HGCN (hyperbolic space) performs competitively on SM and CIDER (hyperspherical space) achieves better results on FI, both demonstrate significant degradation under other OOD types. This underscores the inadequacy of relying solely on a single space to handle the full spectrum of OOD variations. **(4) Diverse nature of distribution deviations**: FI poses the greatest challenge due to the high semantic similarity between ID and OOD samples. SM is particularly difficult for models lacking structural modeling capabilities. In contrast, LL (essentially a large-gap variant of FI) is comparatively easier to detect due to its clearer semantic separation.

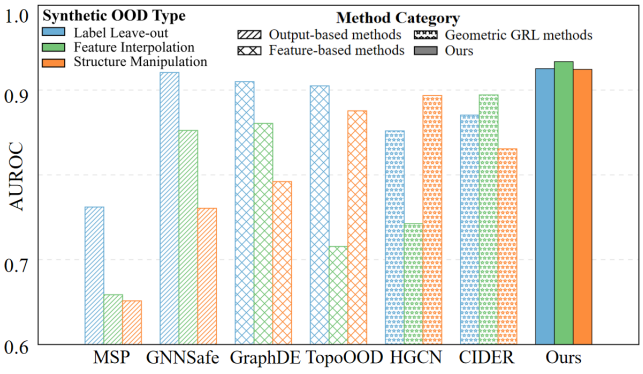

Figure 3: Intuitive illustration of AUROC (↑) performance with three synthetic OOD types in `Ogbn-ArXiv`.

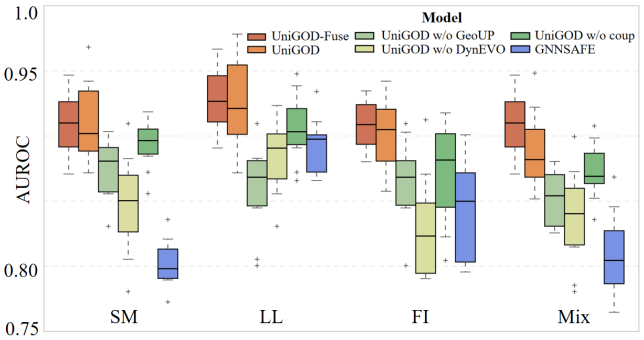

Figure 4: Ablation study of UniGOD and its variants under different OOD types (SM, LL, FI, Mix) in `Ogbn-ArXiv`.

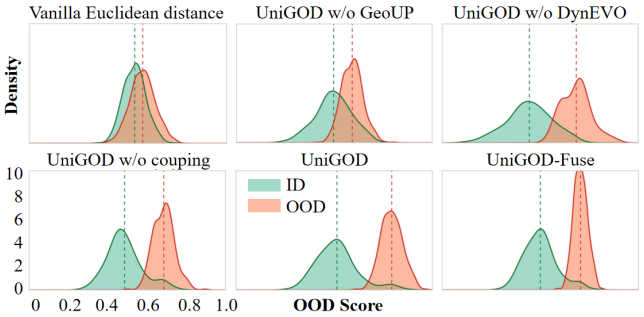

Figure 5: OOD score distribution analysis for UniGOD and its ablated variants under FI OOD in `Ogbn-ArXiv`.

Ablation Study (RQ2)

Fig. 4 & Fig. 5 summarize four key findings: **(1) UniGOD (Eq. 17) *v.s.* UniGOD-Fuse (Eq. 18)**: While their overall performance is comparable, UniGOD-Fuse exhibits greater stability with lower variance, whereas UniGOD shows larger variance but occasionally achieves the exceptionally high performance. Score distributions further indicate that UniGOD yields a larger OOD/ID score gap, while UniGOD-Fuse presents sharper concentration. These results highlight the fusion strategy in leveraging complementary cues across multiple spaces, even when a single optimal space exists. **(2) Impact of GeoUP**: `w/o GeoUP` leads to clear performance degradation, highlighting the role of adaptive curvature learning. The score distribution further reveals reduced OOD/ID separation and higher overlap. This underscores the importance of geometry selection for distinct OOD types. **(3) Impact of DynEVO**: `w/o DynEVO` results in a significant AUROC drop and reduced stability, validating the effectiveness of trajectory uncertainty for OOD/ID separability. **(4) Synergy of geometric-dynamics coupling**: `w/o coupling` causes a consistent performance decline despite both modules being present, underscoring the necessity of joint optimization from the coupling mechanism.

Adaptive Curvature (RQ3)

Fig. 7 demonstrates how representation capacity evolves with curvature, leading to three key insights: **(1) Space preference is OOD-type dependent**: SM peaks at negative cur-

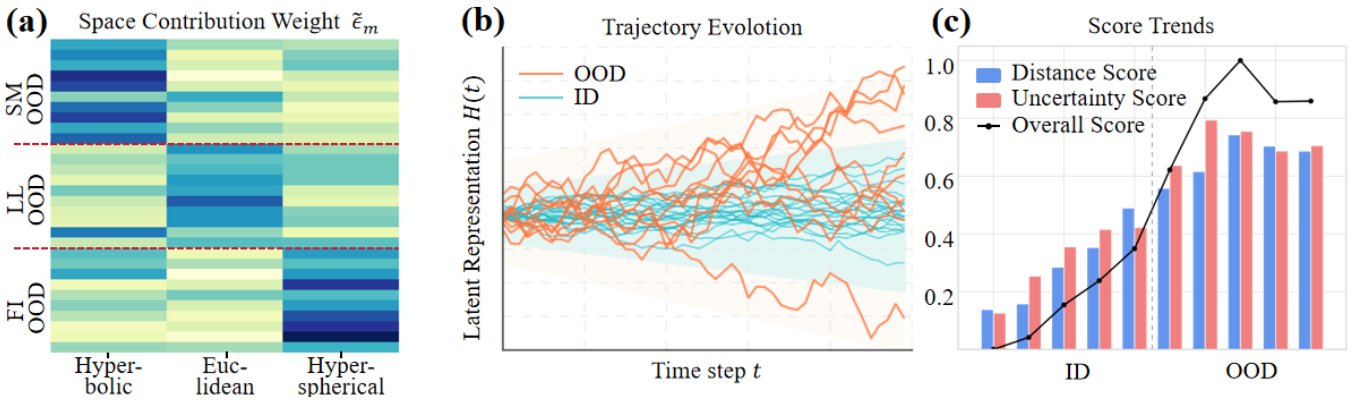

Figure 6: UniGOD offers interpretable insights for its OOD detection decisions.

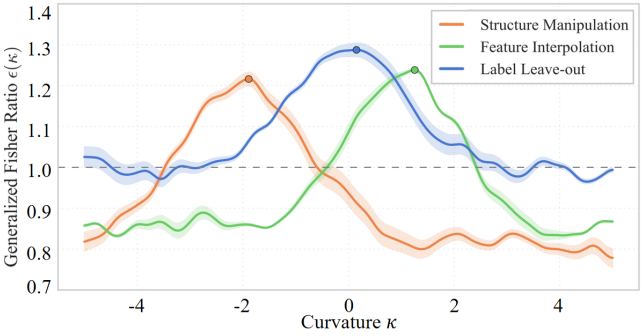

Figure 7: Generalized Fisher Ratio ϵ (↑) under varying curvatures κ on subgraphs of `Ogbn-ArXiv`.

vature ($\kappa = -1.8896$), favoring hyperbolic space, while LL ($\kappa = 0.1505$) and FI ($\kappa = 1.2542$) peak in hyperspherical space, reflecting differing geometric affinities. **(2) Discriminability and curvature sensitivity vary**: LL attains a higher Fisher ratio ($\epsilon = 1.3074$) with minimal curvature ($\kappa = 0.1505$), whereas FI requires stronger deformation ($\kappa = 0.9532$) to reach a lower peak ($\epsilon = 1.2382$), indicating greater intrinsic difficulty. **(3) Curvature impact is non-monotonic**: Moderate positive curvature enhances angular separability, but excessive curvature compresses space, increasing intra-class variance and reducing ϵ. These results underscore the importance of adaptive curvature in modeling diverse OOD shifts.

Parameter Study (RQ4)

Some interesting phenomena are shown in Fig. 8: **(1) Sensitivity to oversmoothing and over-squashing**: Vanilla Euclidean GNNs perform well only at shallow depths ($l < 4$) and exhibit a rapid performance collapse as l increases, due to severe oversmoothing and over-squashing effects. In contrast, our κ-GNN demonstrates exceptional robustness across deeper layers ($l \in [2, 10]$), and can even maintain performance over 20 layers, attributed to curvature-adaptive propagation and geometric regularization. **(2) Effectiveness of temporal parameters**: For SDE temporal step t and step size Δt, the optimal performance is achieved within

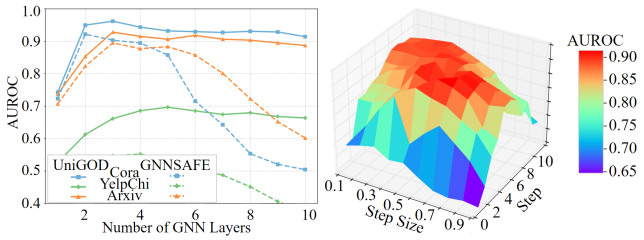

Figure 8: Parameter study for κ-GNN layer depth l (left) and SDE temporal step t and step size Δt (right).

$t \in [3, 6]$ and $\Delta t \in [0.4, 0.7]$, ensuring stable and expressive trajectory evolution to capture uncertainty.

Interpretability Study (RQ5)

Fig. 6 **(a) Geometry-Aware Attribution**: The heatmap visualizes space-specific contributions, uncovering interpretable geometric preferences across OOD types: SM OOD are more salient in hyperbolic space, while FI OOD rely more on hyperspherical space. **(b) Trajectory-Based Uncertainty Disentangles OOD**: Under the SDE-driven dynamics, OOD exhibits greater volatility and deviation over time to form visibly divergent trajectories, offering a clear temporal signal of distributional deviation and enhancing interpretability from a dynamic perspective. **(c) Dual-Factor Scoring Enhances Discriminability**: The multiplicative scoring mechanism combines geometric distance and uncertainty, amplifying their joint contribution. It highlights that OOD nodes are not only distant but also unstable, revealing dual, interpretable evidence for detection.

Conclusion

We revisit the graph OOD detection task under diverse distribution shifts and present UniGOD, the first unified framework that serves as a graph foundation model for this task, exploring adaptive representation across multiple geometric spaces. Beyond achieving SOTA performance, our study provides new insights into the exploration of graph foundation models towards graph OOD detection and graph anomaly detection, offering guidance for future research.

Acknowledgments

This work is supported by the Key Program of the General Technology Joint Fund of the National Natural Science Foundation of China (No. U2336202), the National Natural Science Foundation of China (No.62402491) and the China Postdoctoral Science Foundation (No.2025M771524). We thank all the reviewers for their insights and useful advice.

References

Bergna, R.; Ordoñez, S. C.; Opolka, F.; Lio, P.; and Hernández-Lobato, J. M. 2025. Uncertainty Modeling in Graph Neural Networks via Stochastic Differential Equations. In *Proceedings of the International Conference on Learning Representations (ICLR-25)*.

Cai, T.; Jiang, Y.; Liu, Y.; Li, M.; Huang, C.; and Pan, S. 2025. Out-of-Distribution Detection on Graphs: A Survey. arXiv:2502.08105.

Chami, I.; Ying, Z.; Ré, C.; and Leskovec, J. 2019. Hyperbolic Graph Convolutional Neural Networks. In *Advances in Neural Information Processing Systems (NeurIPS-19)*, volume 32.

Chen, Y.; Luo, Y.; Song, Y.; Dai, P.; and Tang, X., J.and Cao. 2025. Decoupled Graph Energy-based Model for Node Out-of-Distribution Detection on Heterophilic Graphs. In *Proceedings of the International Conference on Learning Representations (ICLR-25)*.

Crane, K.; Livesu, M.; Puppo, E.; and Qin, Y. 2020. A survey of algorithms for geodesic paths and distances. *arXiv preprint arXiv:2007.10430*.

Dong, X.; Zhang, X.; Chen, L.; Yuan, M.; and Wang, S. 2025a. SpaceGNN: Multi-Space Graph Neural Network for Node Anomaly Detection with Extremely Limited Labels. In *Proceedings of the International Conference on Learning Representations (ICLR-25)*.

Dong, Y.; Guo, X.; Xiang, J.; Liu, K.; and Tang, Z. 2021. Hypersphere: An Embedding Method for Knowledge Graph Completion Based on Hypersphere. In *Proceedings of the International Conference on Knowledge Science, Engineering and Management (KSEM-21)*, 517–528.

Dong, Y.; He, R.; Chen, R.; Zhang, W.; Han, Z.; Shi, J.; and Yin, Y. 2025b. G-OSR: A Comprehensive Benchmark for Graph Open-Set Recognition. arXiv:2503.00476v1.

Gong et al., Z. 2024. An Energy-Centric Framework for Category-Free Out-of-Distribution Node Detection in Graphs. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD-24)*, 908–919.

Grover, A.; and Leskovec, J. 2016. Node2vec: Scalable Feature Learning for Networks. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD-16)*, 855–864.

Hendrycks et al., D. 2017. A Baseline for Detecting Misclassified and Out-of-Distribution Examples in Neural Networks. In *Proceedings of the International Conference on Learning Representations (ICLR-17)*.

Kipf, T. N.; and Welling, M. 2017. Semi-Supervised Classification with Graph Convolutional Networks. In *Proceedings of the International Conference on Learning Representations (ICLR-17)*.

Li, Z.; Ao, Y.; and He, J. 2024. SpherE: Expressive and Interpretable Knowledge Graph Embedding for Set Retrieval. In *Proceedings of the ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR-24)*, 2629–2634.

Li, Z.; Wu, Q.; Nie, F.; and Yan, J. 2022. Graphde: A generative framework for debiased learning and out-of-distribution detection on graphs. *Advances in Neural Information Processing Systems (NeurIPS-22)*, 35: 30277–30290.

Liang et al., S. 2018. Enhancing the Reliability of Out-of-Distribution Image Detection in Neural Networks. In *Proceedings of the International Conference on Learning Representations (ICLR-18)*.

Liu, Y.; Ding, K.; Liu, H.; and Pan, S. 2023. Good-d: On unsupervised graph out-of-distribution detection. In *Proceedings of the ACM international conference on web search and data mining (WSDM-23)*, 339–347.

Look, A.; Kandemir, M.; Rakitsch, B.; and Peters, J. 2022. A deterministic approximation to neural SDEs. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI-23)*, 45(4): 4023–4037.

Ming, Y.; Sun, Y.; Dia, O.; and Li, Y. 2023. How to Exploit Hyperspherical Embeddings for Out-of-Distribution Detection? In *Proceedings of the International Conference on Learning Representations (ICLR-2023)*.

Nickel, M.; and Kiela, D. 2017. Poincaré Embeddings for Learning Hierarchical Representations. In *Advances in Neural Information Processing Systems (NIPS-17)*.

Nickel, M.; and Kiela, D. 2018. Learning Continuous Hierarchies in the Lorentz Model of Hyperbolic Geometry. In *Proceedings of the International Conference on Machine Learning (ICML-18)*, 3779–3788.

Perozzi, B.; Al-Rfou, R.; and Skiena, S. 2014. DeepWalk: Online Learning of Social Representations. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD-14)*, 701–710.

Rusch, T. K.; Bronstein, M. M.; and Mishra, S. 2023. A Survey on Oversmoothing in Graph Neural Networks. arXiv:2303.10993.

Song, K.; Solozabal, R.; Hao, L.; Ren, L.; Abdar, M.; Li, Q.; Karray, F.; and Takac, M. 2024. Enhance Hyperbolic Representation Learning via Second-order Pooling. arXiv:2410.22026.

Sun, L.; Ye, J.; Peng, H.; Wang, F.; and Yu, P. S. 2023. Self-supervised continual graph learning in adaptive riemannian spaces. In *Proceedings of the AAAI conference on artificial intelligence (AAAI-23)*, 4633–4642.

Sun, N.; Lin, X.; Zhou, Z.; Shang, Y.; Cheng, Z.; and Cao, Y. 2025. Evidential Spectrum-Aware Contrastive Learning for OOD Detection in Dynamic Graphs. In *Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-25)*.

Wang, P.; Wu, D.; Chen, C.; Liu, K.; Fu, Y.; Huang, J.; Zhou, Y.; Zhan, J.; and Hua, X. 2024. Deep Adaptive Graph Clustering via von Mises-Fisher Distributions. *ACM Transactions on the Web*, 18(2): 1–21.

Wu, Q.; Chen, Y.; Yang, C.; and Yan, J. 2023. Energy-Based Out-of-Distribution Detection for Graph Neural Networks. In *Proceedings of the International Conference on Learning Representations (ICLR-23)*.

Yang, M.; Zhou, M.; Li, Z.; Liu, J.; Pan, L.; Xiong, H.; and King, I. 2022. Hyperbolic Graph Neural Networks: A Review of Methods and Applications. arXiv:2202.13852.

You, S.; Zhang, M.; Liu, J.; Ma, T.; Shen, Y.; and Tang, J. 2022. Graph Out-of-Distribution Detection Goes Neighborhood Shaping. In *Advances in Neural Information Processing Systems (NeurIPS-22)*.

Zhang, G.; Yang, Z.; Wu, J.; Jiao, P.; and Yang, J. 2025. Enhancing Graph Neural Networks for Out-of-Distribution Graph Detection. *IEEE Transactions on Neural Networks and Learning Systems (TNNLS-25)*.

Zhang, K.; Liu, S.; Wang, S.; Shi, W.; Chen, C.; Li, P.; Li, S.; Li, J.; and Ding, K. 2024. A Survey of Deep Graph Learning under Distribution Shifts: from Graph Out-of-Distribution Generalization to Adaptation. arXiv:2410.19265.

Zhao, X.; Chen, F.; Hu, s.; and Cho, J. H. 2020. Uncertainty aware semi-supervised learning on graph data. In *Advances in neural information processing systems (NeurIPS-20)*, 12827–12836.

Zheng, W.; Zhang, G.; Zhao, X.; Feng, Z.; Song, L.; and Kou, H. 2025. Hyperbolic Graph Wavelet Neural Network. *Tsinghua Science and Technology*, 30(4): 1511–1525.

Zhu, S.; Pan, S.; Zhou, C.; Wu, J.; Cao, Y.; and Wang, B. 2020. Graph Geometry Interaction Learning. In *Advances in Neural Information Processing Systems (NeurIPS-20)*, volume 33, 7548–7588.